

# MaRVIN: a distributed platform for massive RDF inferencing

George Anadiotis, Spyros Kotoulas, Eyal Oren, Ronny Siebes, Frank van Harmelen  
Niels Drost, Roelof Kemp, Jason Maassen, Frank J. Seinstra, Henri E. Bal  
Vrije Universiteit Amsterdam

MaRVIN is:

a platform for distributed RDF(S) reasoning  
a platform for processing lots of RDF data

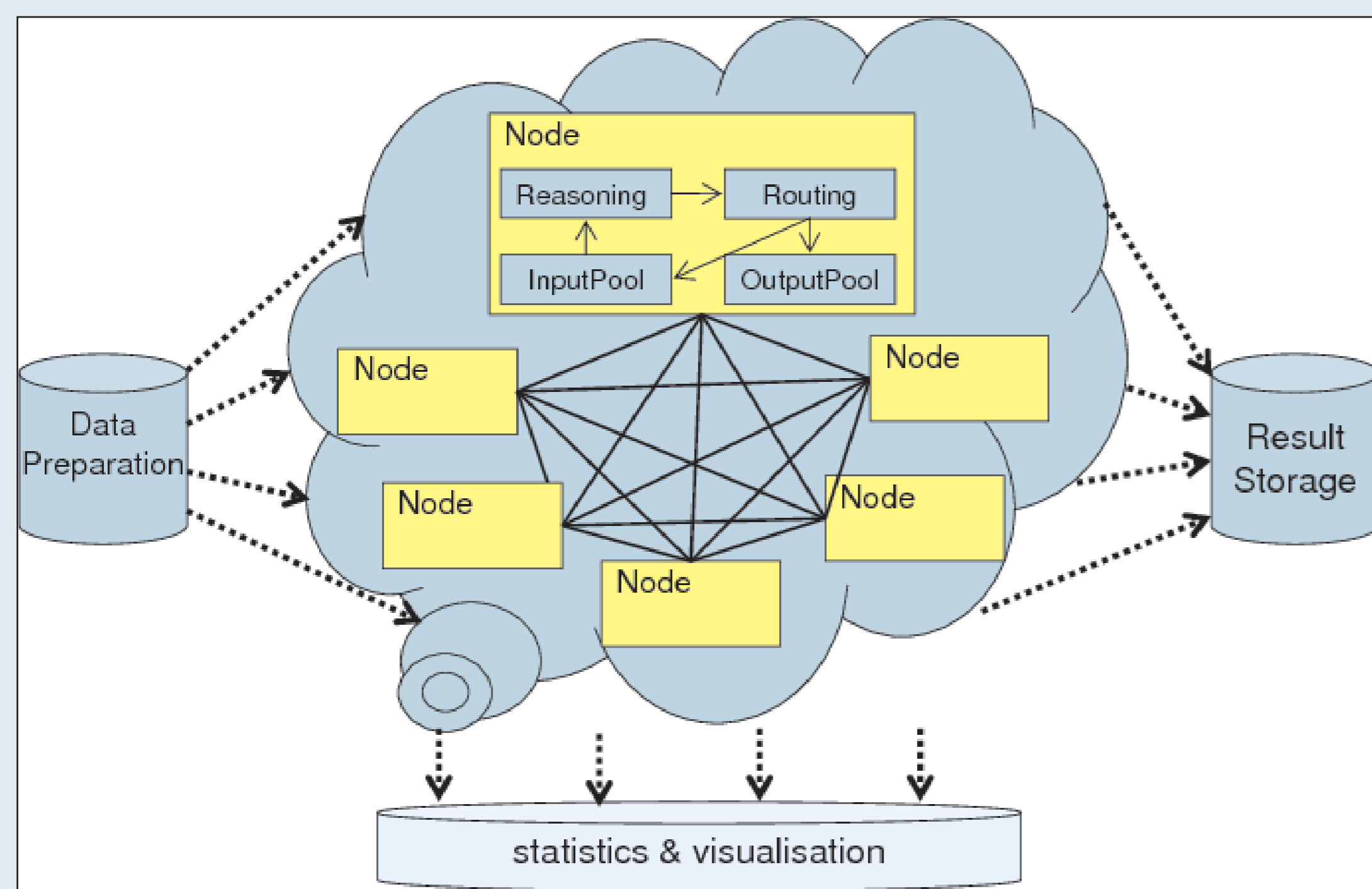
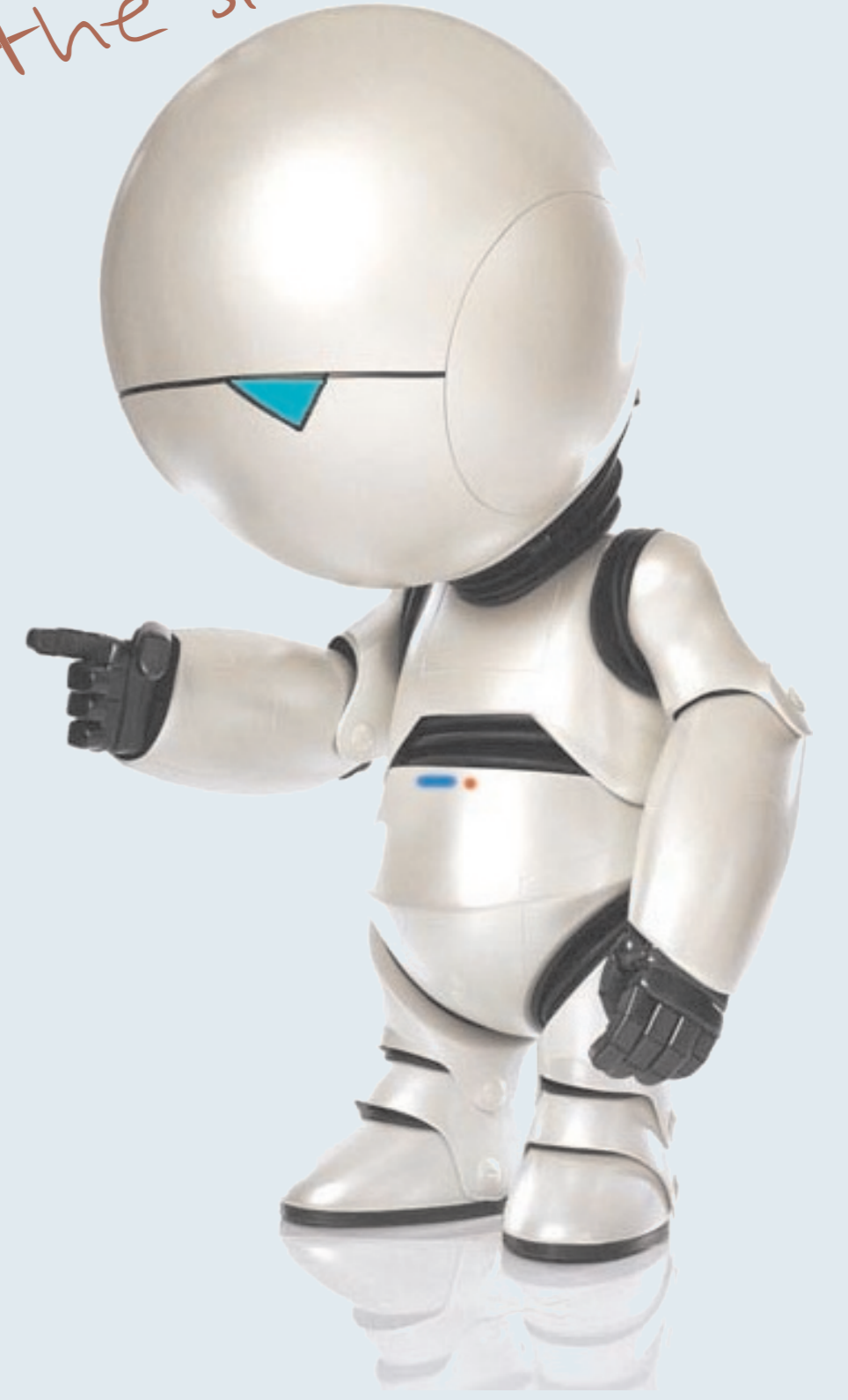
MaRVIN scales by:

distributing computation over many nodes  
approximate (sound but incomplete) reasoning  
anytime convergence (more complete over time)

MaRVIN runs on:

in principle: any grid, using Ibis middleware  
currently: the DAS-3 distributed supercomputer (300 nodes)  
soon: a wide-area a peer-to-peer network

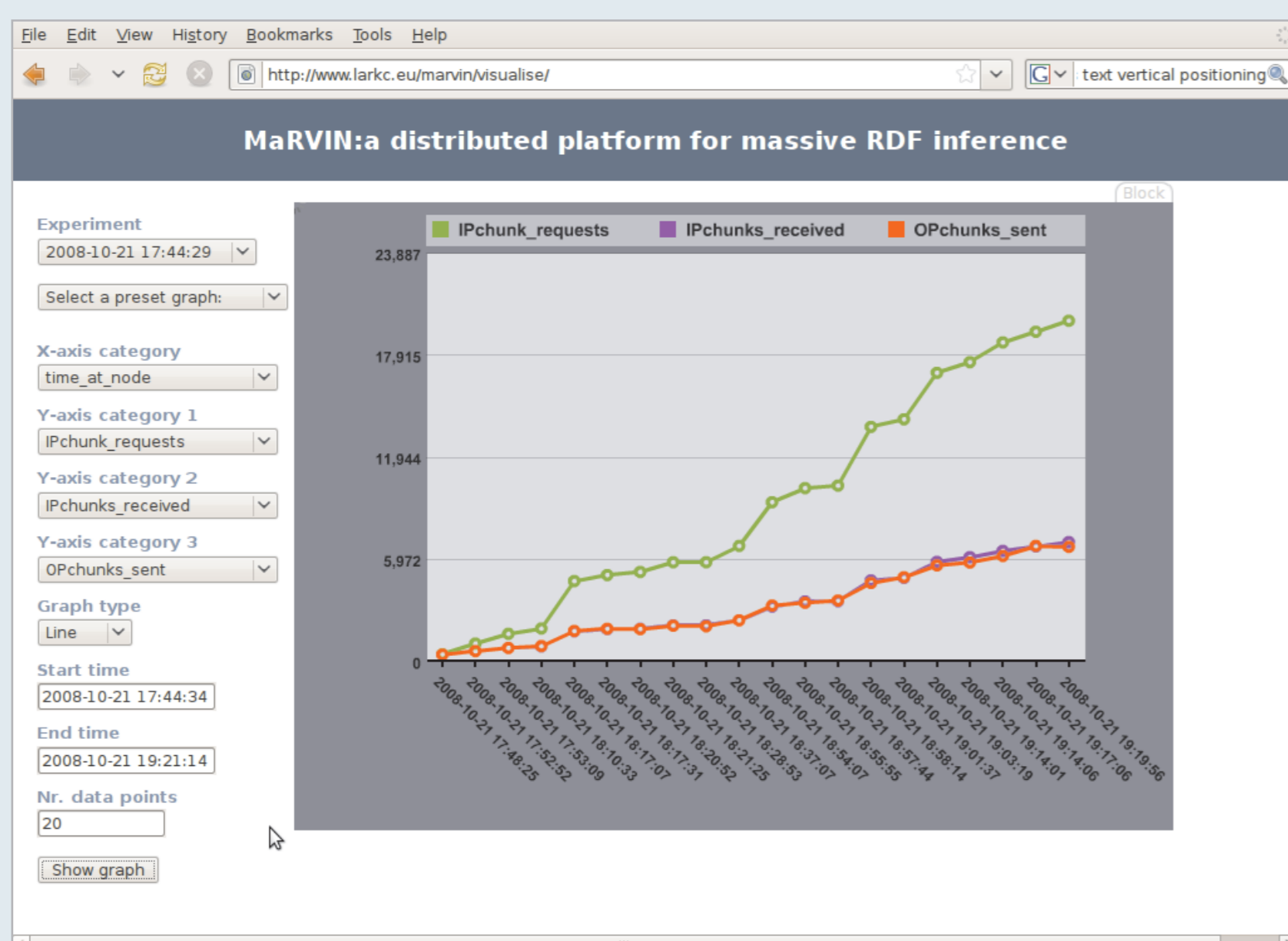
"a brain the size of a planet"



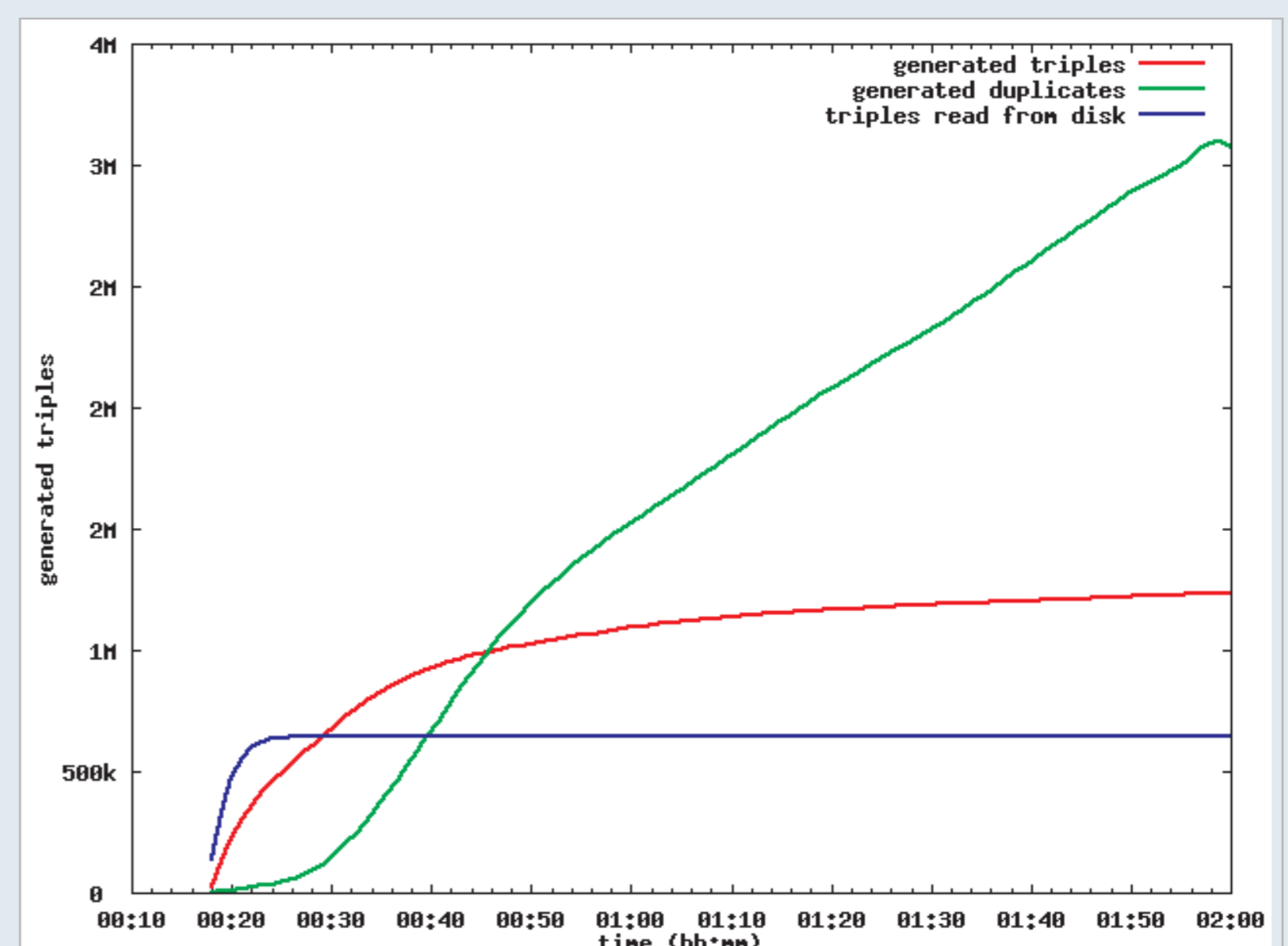
Main loop: divide-conquer-and-swap

1. *divide*: split input data in chunks
2. *conquer*: each node:  
reads some chunks,  
computes closure.
3. *swap*: each node:  
removes all triples:  
sends some to central storage,  
sends other to some peer

repeat 2-3 ad infinitum



real-time logging, visualisation, analysis



anytime incremental results

contact: Eyal Oren  
<http://larkc.eu/marvin>

